Chapter 18
Challenges in Understanding Visual Shape Perception and Representation: Bridging Subsymbolic and Symbolic Coding

Philip J. Kellman, Patrick Garrigan, and Gennady Erlikhman

18.1 Introduction

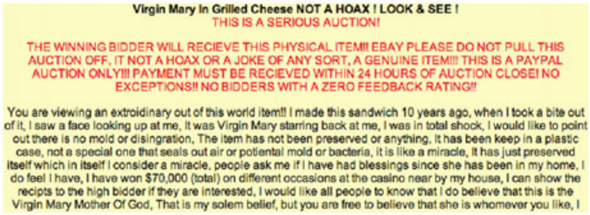
Our everyday perceptual experience is of a world populated by objects and surfaces arrayed in space, as well as of events that produce changes in these arrangements over time. Successful perception, thought and action depend on processes that produce accurate descriptions of these objects and events. Often, object contours are only partially visible as we move or as they move around us. Nevertheless, we experience a unified, stable world: the squirrel running through the tree branches appears as a single animal, not as dissociated squirrel-bits, and the house seen through the slats of a fence is one house, not a collection of independent house fragments. These perceptual outcomes depend on a number of segmentation, grouping, and interpolation processes, which, taken together, perform some of the most crucial and remarkable tasks in allowing us to perceive the world visually. They also pose some of the greatest challenges in understanding the underlying processes and mechanisms of vision.

Researchers in the past several decades have made considerable progress on a number of important components of these perceptual capabilities. Much is known about early cortical processing of visual information. At a more abstract level, experimental data and computational models have revealed a great deal about contour, object, and shape perception. Neurophysiological and imaging methods have provided evidence for functional specificity in areas of cortex for animate and inanimate objects, tools, faces, and places. However, between the initial encodings by spatially localized units and higher level descriptions of contours, surfaces, objects and their properties lies a considerable gap. To use a chess analogy, we do not understand

P.J. Kellman (✉) · G. Erlikhman
Department of Psychology, University of California, Los Angeles, USA
e-mail: kellman@cognet.ucla.edu

P. Garrigan
Department of Psychology, St. Joseph's University, Philadelphia, PA, USA

S.J. Dickinson, Z. Pizlo (eds.), *Shape Perception in Human and Computer Vision*, Advances in Computer Vision and Pattern Recognition,
DOI 10.1007/978-1-4471-5195-1_18, © Springer-Verlag London 2013



Virgin Mary In Grilled Cheese NOT A HOAX ! LOOK & SEE !
THIS IS A SERIOUS AUCTION!

THE WINNING BIDDER WILL RECIEVE THIS PHYSICAL ITEM!! EBAY PLEASE DO NOT PULL THIS AUCTION OFF, IT NOT A HOAX OR A JOKE OF ANY SORT, A GENUINE ITEM!!! THIS IS A PAYPAL AUCTION ONLY!!! PAYMENT MUST BE RECIEVED WITHIN 24 HOURS OF AUCTION CLOSE! NO EXCEPTIONS!! NO BIDDERS WITH A ZERO FEEDBACK RATING!!

You are viewing an extroidinary out of this world item!! I made this sandwich 10 years ago, when I took a bite out of it, I saw a face looking up at me, It was Virgin Mary starring back at me, I was in total shock, I would like to point out there is no mold or disingration, The item has not been preserved or anything, It has been keep in a plastic case, not a special one that seals out air or potiental mold or bacteria, it is like a miracle, It has just preserved itself which in itself I consider a miracle, people ask me if I have had blessings since she has been in my home, I do feel I have, I have won $70,000 (total) on different occasions at the casino near by my house, I can show the recipts to the high bidder if they are interested, I would like all people to know that I do believe that this is the Virgin Mary Mother Of God, That is my solem belief, but you are free to believe that she is whomever you like, I

A B

Fig. 18.1 Spiritual, culinary, and commercial aspects of shape perception. (**A**) This 10-year-old, partially eaten cheese sandwich sold for $28,000 on Ebay; the owner claimed to see the face of the Virgin Mary in it. (**B**) Description from the Ebay ad. (See text.) (From http://www.slate.com/articles/news_and_politics/explainer/2004/11/the_28k_sandwich_that_grew_no_mold.html)

much about the "middle game." The study of shape perception and representation is important in its own right but also because it gives us a sharp focus on some of the biggest unsolved general issues in the computational and neural understanding of perception.

Early cortical encodings (e.g., responses of neural units in V1) are spatially local, retinally specific, and modulated by oriented contrast. The functionally important outputs of perceiving are constancy-based descriptions of bounded objects, their contours, surfaces, and shapes, and their arrangements in space. Our goal in this chapter is to shed light on shape perception, but also to use it as a vehicle to focus on major issues that must be addressed in order to understand how early visual processes connect to high-level representations. We describe (1) the dependence of shape perception on segmentation and grouping processes, and (2) properties that (some) shape representations must have and how they might be assembled from lower level encodings. In both discussions, we end with thoughts and efforts on a crucial frontier of work in these areas, which we might call "modeling the middle."

18.2 Some Useful Examples

To begin, we offer two demonstrations that illustrate the flexible and abstract nature of shape representations and the important issue of what gets assigned a shape representation.

What is shown in Fig. 18.1? It is perhaps the most famous cheese sandwich in history. The story, in the owner's own words in an E-bay advertisement, is given in Fig. 18.1B. The image and description may relate to several different scientific mysteries. Given that this cheese sandwich has had "no disingration [sic]" in 10 years, one of the mysteries is, obviously: What are they in putting in the bread?!

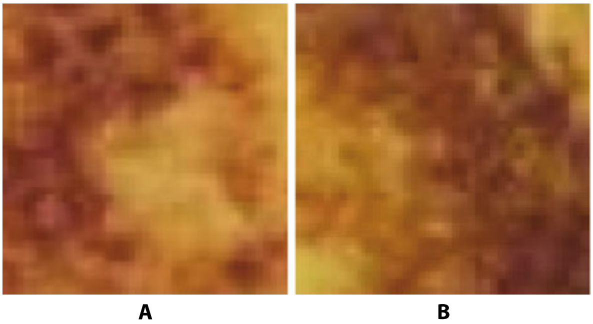


Fig. 18.2 Which region was part of the display in Fig. 18.1A? The difficulty of answering provides a simple demonstration that we encode simplified and abstract descriptions of displays, not pixel maps or records of feature activations

For us, the more important fact is that humans spontaneously see a face in the toast. This reveals more than one interesting fact about shape representations. A key observation is that such representations are flexible enough to be matched to new input that is markedly different from previously experienced faces. Putting aside whether the image is in fact the Virgin Mary, or bears, as others have suggested, a resemblance to Greta Garbo or a young Shirley Temple, the striking fact is that any recognition here is not a match at the pixel level. Presumably, if you have seen Garbo before, she was not impersonating a cheese sandwich. More formally, our encoding of this display is both much less and much more than a literal copy of the stimulus. Consider Fig. 18.2. Suppose we tell you that one of the panels shows a region of Fig. 18.1A. Without looking at Fig. 18.1A, which is it—Fig. 18.2A or 18.2B? We doubt anyone can answer correctly with confidence. Now, compare the regions to Fig. 18.1A. With serious effort, you can see that the region in Fig. 18.2A matches an area near the left eye, and the region in Fig. 18.2B also matches, around the right eye. Even with all images visible, verification of the match is an effortful task. This kind of demonstration, and many ordinary observations, indicate that we preserve very little of the point-by-point stimulus in encoding shape information.

This should not be seen as a shortcoming of our visual processing. The ability to see a face here and to detect similarities to previously seen faces implicates extraction of relevant structure while ignoring irrelevant variation. The structure extracted must be encoded in some abstract form sufficient to trigger activation of previously encoded structure. Two other notable points here are that we are able to see a face while still encoding the entire object as a cheese sandwich, and that we spontaneously see the face despite the low expectations (and low prior odds, in Bayesian frameworks) for seeing faces in partially eaten cheese sandwiches. Most important, however, is the suggestion of an abstract, flexible representational format that allows matching of selected structure to categorical shape information encoded or formed earlier.

Our second demonstration leads more directly into the connection between shape perception and visual segmentation and grouping processes. Glance at the picture in

Fig. 18.3 Illustration of the dependence of shape descriptions on object formation. (See text.) (Reprinted with permission from fotosearch.com)

Fig. 18.3 and then cover it up. Looking at Fig. 18.4, which panel, A or B, shows a shape that was present in the original figure? The question is difficult to answer.

Now uncover Fig. 18.3. Both regions turn out to be part of the picture. The region in Fig. 18.4A is part of the cow's head, and the shape in Fig. 18.4B is part of the fence post. Both of the regions shown in Fig. 18.4 are fairly well delineated by contrast boundaries in the image.[1] Naively, we might expect that bounded regions in the visual input comprise the objects to which we assign shape descriptions. Examples such as these demonstrate that such an expectation is often incorrect.

We assign shape descriptions to *objects*. The detection of objects in a visual scene, if it is to correspond to actual physical objects in the world, must overcome a number of obstacles. Perhaps most important is occlusion. A single object in the world may project to the retinae of our eyes in multiple, spatially separated regions, as illustrated by the cow in Fig. 18.3. A single object may have a variety of colors, such that lightness and color boundaries are incomplete indicators of object boundaries. Fortunately, our visual processing includes sophisticated mechanisms

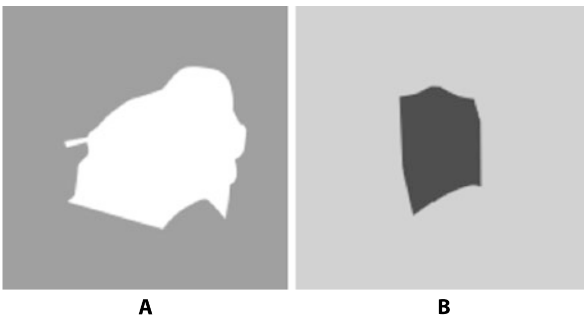


Fig. 18.4 Which region is part of the display in Fig. 18.3? (See text)

[1] These are mostly, but not fully, delineated by contrast boundaries. The difficulty in using contrast boundaries alone to find the functionally important shapes in the environment is another important aspect of the relation between processes that accomplish segmentation and shape representation.

for perceiving coherent objects from information that is fragmentary in space (and also in time, although we do not consider spatiotemporal fragmentation here; see [37] for relevant research and [32] for a review).

Shape perception in biological vision, then, means something more than finding regions of roughly homogeneous lightness and/or color. Rather, shape encoding appears to be reserved for functional units delivered by segmentation and grouping processes. In the next section, we describe processes of interpolation that connect visible regions across gaps to furnish the objects that receive shape descriptions. Introducing these will show their relevance to shape perception and also highlight the issue of "modeling the middle" in vision science.

18.3 Interpolation Processes Underlying Object Perception

From the perspective of an organism that needs to see, the projection of objects and scenes in the world onto the sensitive surfaces of our eyes is beset by several chronic problems. The world has three spatial dimensions, but information is lost as it is projected onto the essentially two-dimensional surface of each retina. Light moves in straight lines, and objects are usually opaque; these facts dictate that in ordinary environments nearer objects will often partly occlude farther ones, meaning the projections of farther objects will be interrupted by the projections of nearer ones. Commonly, a single object may project to multiple, spatially separated retinal regions.

When motion of objects or observers is involved, these patterns of occlusion become more complex, changing over time. Different parts of a single object may be visible at different times, while some parts of objects may never project to the eyes at all. Such problems of occlusion are not exclusively products of modern, cluttered environments; some of the richest and most complex patterns of occlusion occur when we view objects and scenes through foliage, a situation that has likely been important in human behavior over evolutionary time.

Perhaps these enduring constraints on seeing are responsible for the sophisticated and elegant visual processes that serve to overcome occlusion. The human visual system possesses remarkable mechanisms for recovering coherent objects and surface representations from fragmentary input. Specifically, object and surface perception depends on interpolation processes that overcome gaps in contours and surfaces in 2-D, 3-D, and spatiotemporal displays. Recent research suggests that the mechanisms for doing so are deeply related in that they exploit common geometric regularities.

18.4 Contour and Surface Processes

Evidence suggests that there are two kinds of mechanisms for connecting visible areas across gaps: contour and surface interpolation. These processes can be distinguished because they operate in different circumstances and depend on different

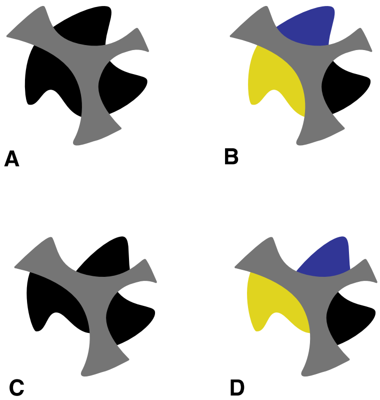


Fig. 18.5 Contour and surface interpolation. (**A**) Both contour and surface interpolation processes contribute to perceived unity of the three black regions behind the gray occluder. (**B**) Contour interpolation alone. (**C**) Surface interpolation alone. (**D**) Both contour and surface interpolation have been disrupted, causing the *blue*, *yellow*, *and black regions* to appear as three separate objects. (See text)

variables. Contour interpolation depends on geometric relations of visible contour segments that lead into contour junctions. These geometric constraints have been most frequently studied in 2-D displays, but they have been shown to govern contour interpolation in 3-D scenes as well [33]. Surface interpolation in 2-D displays can occur in the absence of contour segments or junctions; it depends on the similarity of lightness, color, and/or texture of visible surface patches. In 3-D scenes, it also depends on the orientations and positions of visible fragments [9]. For simplicity, we describe interpolation processes in 2-D displays here.

Figure 18.5 illustrates distinguishable contour and surface interpolation processes, as well as some of their interactions. In Fig. 18.5A, the three black regions appear as one object connecting behind the gray occluder. Both the contour relationships of the black regions and their surface similarity contribute to this percept. In Fig. 18.5B, the surface colors of the visible regions have been altered to block surface interpolation. However, the relations of the contours still engage contour interpolation, leading to an impression of a unified object despite the color differences. Figure 18.5C shows the converse arrangement. Here, the geometry of *contour relatability* (see below) has been disrupted blocking contour interpolation. Due to surface interpolation (included by the matching surface color of the fragments), however, there is still some impression that the three fragments connect behind the occluder. Finally, Fig. 18.5D shows both contour and surface interpolation disrupted. Here, the blue, yellow, and black regions appear as three separate objects.

Figure 18.6 further illustrates the action of surface interpolation. Surface interpolation in Fig. 18.6B causes the same black objects that appear separate in Fig. 18.6A to appear connected. Surface interpolation also causes the circle within the gray area to appear as a hole, rather than a spot on top of a surface [54]. In this display, con-

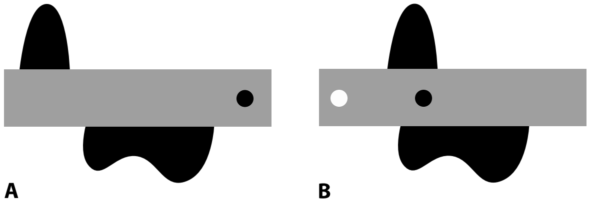


Fig. 18.6 Rules of surface interpolation under occlusion. Contour relations in both displays are arranged so as not to produce contour interpolation behind the occluder. (**A**) The three black regions appear as separate objects; the *circle on the right* appears as a spot on top of the gray background. (**B**) The three black regions have been positioned so that surface spreading within extended tangents of edge orientations at points of occlusion allows areas to connect behind the occluder. A bloblike single object, whose contours behind the occluder are vague, is perceived. The *black circle* is now seen as a *hole* in the occluder. The *white circle* also illustrates surface spreading; it appears as a hole through which the white background is seen

tour interpolation is blocked due to misalignment of the edges. It has been shown that the surface interpolation process under occlusion integrates areas of similar surface quality (1) when they fall within edges connected by contour interpolation, (2) when they fall within the extended tangents of nonrelatable edges, or (3) when the fall within a fully surrounding area (as in the case of the white dot in Fig. 18.6B) [55]. Whereas contour interpolation processes are relatively insensitive to relations of lightness or color, the surface process depends crucially on these.

18.5 Contour Interpolation

Central to establishing perceived shape is the process of contour interpolation, which unifies visible regions across gaps (for reviews, see [31, 33]). Perhaps the most basic question in understanding visual object and surface formation is what stimulus relationships cause it to occur. This question is fundamental because it allows us to understand the nature of visual interpolation. For contour interpolation, certain relations of visible contours lead the visual system to fill in connections between visible regions whereas other contour relationships do not. Discovering the geometric relations and related stimulus conditions that lead to object formation is analogous to understanding the grammar of a language (e.g., what constitutes a well-formed sentence). This level of understanding is also most crucial for appreciating the deepest links between the physical world and our mental representations of it. While these efforts are at first descriptive, as unifying principles are revealed, they allow us to relate the information used by the visual system to the physical laws governing how objects and surfaces project to the eyes, in the form of deep constraints about the way the world works (e.g., [16, 36]) or as natural scene statistics (e.g., [14]).

18.6 Triggering Contour Interpolation

A general fact about contour interpolation is that interpolated contours begin and end at junctions or corners in visible contours (tangent discontinuities). These are locations at which contours have no unique orientation [44, 46]. Most typically in vision, they are intersections of two oriented contours, such as "T" junctions that form when the boundary of an occluding object interrupts that of an occluded object. Whereas a zero-order discontinuity would be a spatial gap in a contour, a first-order or tangent discontinuity is a point at which the direction of the contour changes abruptly. Besides first-order discontinuities, some have suggested that second-order discontinuities (as where a straight segment joins a constant curvature segment, with the slopes matching at the join point) might also play a role in triggering interpolation ([2–4, 46]; for discussion see [33]). The importance of tangent discontinuities in visual processes coping with occlusion stems from an ecological invariant: Shipley and Kellman [46] observed that in general, interpolated contours begin and end at tangent discontinuities and showed that their removal eliminated or markedly reduced contour interpolation. In the patterns that induce illusory contour formation, "L" junctions, rather than T junctions, are most common. In these displays, the presence or absence of tangent discontinuities can be manipulated by rounding the corners of inducing elements, a manipulation that experimental evidence shows reduces or eliminates contour interpolation (e.g., [3, 33, 37, 46]).

18.7 Contour Relatability

What determines which visible contour fragments get connected to form objects? Although tangent discontinuities are ordinarily necessary conditions for contour interpolation, they are not sufficient. After all, many corners in images are corners of objects, not points at which some contour passes behind an intervening surface (or in front, as in illusory contours).

Empirical research shows that contour interpolation depends crucially on geometric relations of visible contour fragments, specifically the relative positions and orientations of pairs of edges leading into points of tangent discontinuity [11, 26, 29–31, 33, 37, 44, 46]. These relations have been described formally in terms of *contour relatability* [29, 49]. Relatability is a mathematical notion that defines a categorical distinction between edges that can connect by interpolation and those that cannot (see [29]). The key idea in contour relatability is smoothness (e.g., interpolated contours are differentiable at least once), but it also incorporates monotonicity (interpolated contours bend in only one direction) and a 90° limit (interpolated contours bend through no more than 90°). Figure 18.7 shows a construction that is useful in defining contour relatability. Formally, if E_1 and E_2 are surface edges, and R and r are perpendicular to these edges at points of tangent discontinuity, then E_1 and E_2 are relatable if and only if:

$$0 \leq R\cos\theta \leq r.$$

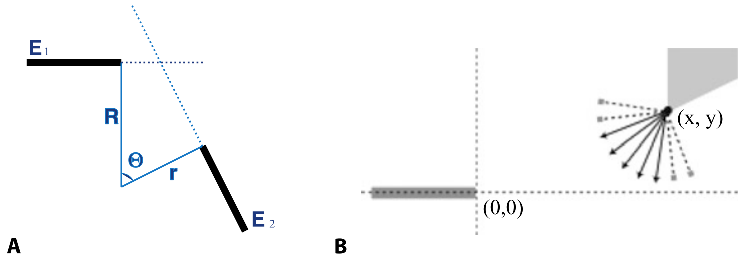


Fig. 18.7 Contour relatability. Contour relatability describes formally a categorical distinction between edges that can be connected by visual interpolation and those that cannot. (**A**) Geometric construction defining contour relatability (see text). (**B**) Alternative expression of relatability. Given one visible contour fragment terminating in a contour junction at (0, 0) and having orientation 0 deg, those orientations Θ that satisfy the equation $\tan^{-1}(y/x) \leq \Theta \leq \pi/2$ are relatable. In the diagram, these are shown with *solid lines*, whereas nonrelatable orientations are shown with *dotted lines*. (Adapted from [23]. A unified model of illusory and occluded contour interpolation. *Vision Research*, *50*, 284–299. Reprinted with permission)

Although the precise shape of interpolated contours is a matter of some disagreement, there are two properties of relatability that cohere naturally with a particular class of contour shapes. First, it can be shown that interpolated edges meeting the relatability criteria can always be comprised of one constant curvature segment and one zero curvature segment. Second, it appears that this shape of interpolated edges has the property of being a minimum curvature solution in that it has lowest maximum curvature: any other first-order continuous curve will have at least one point of greater curvature [29]. This is a slightly different minimum curvature notion than minimum energy.

Relatability is primarily a categorical distinction, indicating which edges can be connected by contour interpolation. Object perception often involves a discrete determination of whether two visible fragments are or are not part of the same object. Figure 18.8 shows examples of relatable and nonrelatable edges, in both perception of partly occluded objects and perception of illusory objects. Complete objects are formed in the top row but not in the bottom row. In general, object formation has profound effects on further processing, such as generation of a representation of missing areas, generation of an overall shape description, and comparison with items or categories in memory. Research indicates that the representation of visual areas as part of a single object or different objects has many important effects on information processing [6, 33, 56].

Relatability is a mathematical formulation that accounts for empirical findings on the geometric relations that support contour interpolation. It incorporates several separable claims, all of which have received substantial confirmation in empirical research. These include the requirements that the edge fragments that participate in interpolation are those terminating in tangent discontinuities [18, 33, 44, 46], the requirements that interpolated edges have orientations matching their inducing edges at the points of tangent discontinuity, are smooth (differentiable at least once), and monotonic (i.e., they do not doubly inflect) [10, 11, 29, 33, 37, 48]. Most evidence

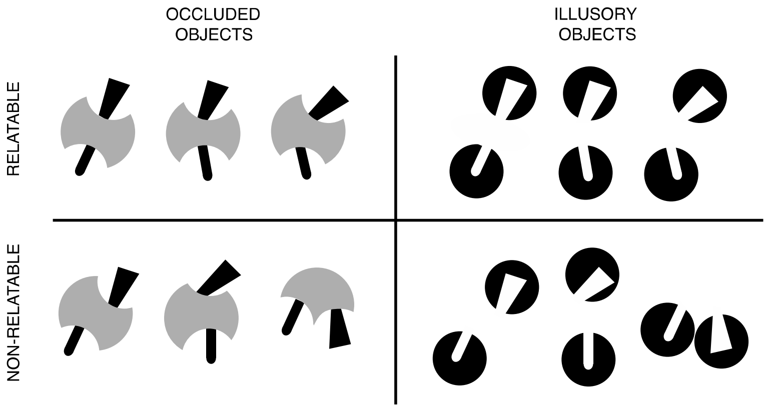


Fig. 18.8 Examples of relatable and nonrelatable contours. (See text)

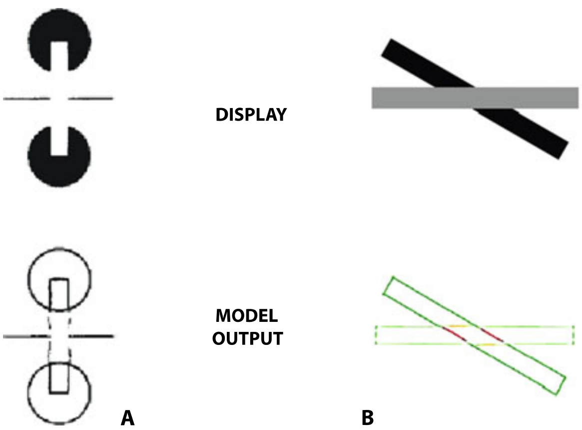


Fig. 18.9 Outputs of neurally plausible models of contour interpolation. The *top row* shows the raw image given to each model, and the *bottom row* shows the output (real and interpolated edges). (**A**) Output of the Heitger et al. [21] model for illusory contour display. (**B**) Output of Kalar et al. [23] model for an occlusion display. (Adapted from Kalar et al. [23], *Vision Research*. Reprinted with permission)

also supports the idea that interpolation is weak or absent for Θ greater than 90 deg in Fig. 18.9B [10, 11, 14, 26, 33, 49], although data also suggests that the cutoff may not be abrupt [11, 19]. Although discrete classification of visible areas as connected or not is important, there is also evidence that quantitative variation exists within the category of relatable edges [5, 10, 29, 47–49]. Singh and Hoffman [49] proposed an expression for quantitative decline of relatability with angular change.

Some work based on scene statistics has been interpreted as showing some deviations from the predictions of relatability. Geisler & Perry [13] reported statistics about the probabilities of arbitrary contour segments being connected in a variety of scenes. In the same paper, the authors reported that observers' subjective judgments of contour connectedness conformed reasonably well to the scene statistics. Compared to relatability, the most systematic deviation appeared to be that relatability allows connections between edge fragments of opposite contrast polarity, a phenomenon that has been confirmed experimentally [10, 15, 24, 28], whereas the collected scene statistics indicate that such connections are highly improbable. The authors' data also indicates that an ideal observer using the natural priors they obtained would interpolate only between very nearly collinear edge fragments and primarily those within one deg of separation. These outcomes are surprising, in that they markedly differ from considerable evidence obtained from a variety of paradigms about human contour interpolation [11, 18, 33, 42, 48].

The discrepancies are not difficult to understand, however. The scene statistics gathered by Geisler & Perry [13] involved the probabilities of any arbitrary edge fragments in scenes being connected. A key geometric invariant in contour interpolation is that occlusion produces tangent discontinuities in the optical projection [29] and evidence indicates that this information is influential in contour interpolation (e.g., [44, 46]). Sampling edge fragments terminating in tangent discontinuities would involve a more restricted set of edge pairs and these may produce different scene statistics. The conditional probability of a pair of edge fragments being part of the same contour in the world, given their relative orientation, position, and separation may differ from the conditional probability of a pair of edge fragments being part of the same contour given those spatial relations *and* the fact that each terminates in a contour junction (typically a T junction, for potential cases of amodal completion). The latter seems more relevant to understanding the relations of environmental regularities to contour interpolation. We do not know whether these two conceptually different conditional probabilities would differ in their empirical distributions, but intuitively, the locations, orientations, sizes, etc. of occluders seems unlikely to be uniformly distributed across images.

Also difficult to interpret is the empirical study reported by Geisler & Perry [13], which involved subjective judgments of 7 observers, two of whom were not naïve. Observers were instructed that half of edge pairs presented in the study would be connected. Such instructions seem incompatible with an attempt to assess participants' natural perceptions of whether two edges appear to be connected under occlusion or not. These instructions also did *not* reflect the priors derived from scene statistics, so observers' results were compared to arbitrarily revised scene statistics incorporating a 0.5 prior on edges being connected, a prior that far exceeded the "natural priors" obtained from Geisler & Perry's scene statistics. Unlike many studies that have used objective performance methods [10, 30, 33, 37], the subjective report methods employed by Geisler & Perry [13] in combination with the prompting of participants to judge 50 % of edge pairs as connected make the task fraught with demand characteristics, as well as difficult to relate either to scene statistics or to other data on interpolation performance. One other major difference from both prior research and ordinary perception of natural scenes is that each "edge" presented in the experiment was a tiny Gabor element (with length roughly 6–7 arc min), and pairs of elements had comparatively large separations (occluders had diameters of 40, 80, and 180 arc min). It is known that strength of interpolation between pairs of inducers is a roughly linear function of support ratio [5, 48, 50], defined as the length of interpolated edge as a fraction of total (real plus interpolated) edge length. Relatively little or very weak interpolation would be expected with support ratios ranging from 0.07–0.25, as in this study, and the scene statistics in this study did not incorporate inducing edge lengths in any manner. It would be interesting to study the relationship of contour relatability to richer scene statistics in future research. Existing data support the geometric relations encompassed by relatability as a formal account of human contour interpolation, and the value of this particular geometry might indeed bear close relations to relevant statistical regularities in natural scenes.

18.8 Neural Models of Contour Interpolation

The model of object formation from fragmentary information, as we have sketched it here, as well as in more elaborate treatments [23, 25, 26, 33], assumes that certain kinds of inputs have been identified in prior visual processing. The inputs to contour interpolation, for example, are oriented edges of surface regions. The contour orientations that matter are those leading into tangent discontinuities, which we assume can be located by earlier visual processing. Whether such contour inputs connect depends on geometric relations of their orientations, which we assume are also encoded. Once a contour segment is interpolated, it, along with the physically given parts of the contour, become a continuous contour that closes, defining the boundary of some object. To these closed contour tokens, we assign a perceived shape.[2]

A variety of neural-style models have been proposed as giving the underlying, neurally plausible mechanisms by which the above computations are performed [17, 21, 23, 31]. Therefore, it would seem that the general issues we raised at the start of this paper have been addressed: High-level information processing accounts of visual object completion have been connected to plausible neural mechanisms, providing not necessarily final or correct explanations, but at least explanations that show how we can go from initial registration of visual information to high-level scene descriptions.

This impression, however, would be as illusory as many of the contours perceived in the visual completion literature. Although neural-style models of visual completion exist, in general, they illustrate, rather than solve, the problem of bridging low-level visual coding and higher level, symbolic representations. Figure 18.9 helps to illustrate the issues.

In the figure are displays presented to two contour interpolation models (Fig. 18.9A) as well as the outputs of those models (Fig. 18.9B). The display and output on the left are from Heitger et al. [21] and those on the right are from Kalar et al. [23]. It is evident that the models fill in illusory and occluded contours based on the input contours. These models use local oriented edge detectors and grouping

[2]Obviously, this brief description leaves out many additional specifics. For example, our treatment here has been confined to 2-D interpolation and the "object" formed by completing the boundary would be a planar (2-D) object. Consideration of 3-D and spatiotemporal object formation is discussed in more detail elsewhere [33, 37], but the current treatment is sufficient to raise the general issues about modeling that are the focus of this section.

operators that examine the relations of activated units to determine interpolated activation in the space in between. Each pixel in the interpolated area is the output of a grouping operator that was positioned at that location.

These models and others (e.g., [17]) show that an early stage of interpolation can be done by sets of local operators that look at relations of contour activation in nearby regions and produce activation maps for regions in between. The input operators (edge detectors) and grouping operators are consistent with known characteristics of early visual cortical areas [20, 21], and their outputs likely approximate an important early stage in object formation [23]. However, it is crucial to understand what these models do and do not do. Specifically, the models have indicated points of interpolation in areas with no stimulus contrast, but they don't do much else. When we look at the output images, we see complete contours that span between input edges, but the models do not connect the interpolation points into contour tokens, nor do they connect the interpolated and real contours into contour tokens. They also do not certify whether these contours close, assign shapes to either the contour parts or the enclosed regions, determine what the objects are, or indicate which object is closer in the display. For example, in the display in Fig. 18.9B, we see two rectangles, with the gray rectangle partly occluding the black one. Given edge orientations and positions, the computational interpolation model of Kellman & Shipley [29] would interpolate the edges as shown. The model of Kalar et al. [23], intended as a neurally plausible implementation of the Kellman & Shipley model that operates on raw images, produced the image in Fig. 18.9B. The model's output, however, and the predecessor model of Heitger et al. [21], consists of a collection of points of "interpolation activation": it marks where interpolated edges would occur, but it does not produce a representation of connected edges, closed objects or depth relations. The apparent continuity of contours and shapes of closed figures are generated by the viewer when they look at the model's output image. The model itself does not "know" what is connected to what. We might call these "subsymbolic" models. Thus, "local" interpolation models leave a lot of work to be done. They build from the kinds of spatially localized neural units that exist in cortex, but they stop short of giving contour and object descriptions needed for higher level representations. Those descriptions need to be much more abstract, symbolic representations, as we discuss in the next section.

18.9 Shape Perception

Modern work in computational vision has typically addressed shape with a variety of sophisticated mathematical techniques (for a review, see [8]). These techniques allow great precision. For contour shape, having even a few data points allows, for example, polynomial approximation that specifies all of the contour's derivatives at all points. Yet neither these computational techniques nor neurophysiological data about the functions of cortical neurons has yet produced a real understanding of shape perception in biological vision.

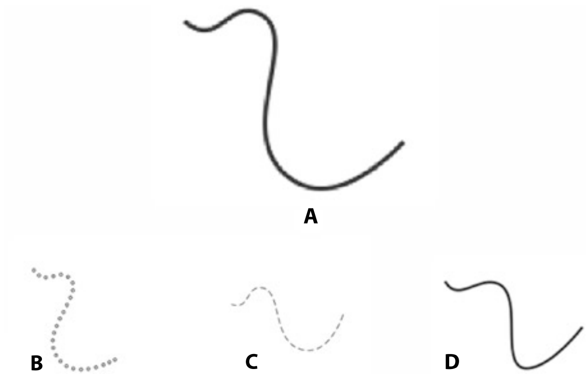


Fig. 18.10 Examples of shape invariance. Which of the figures in **B**, **C**, or **D**, has the same shape as the figure in **A**? (See text)

To understand some of the key issues in shape perception, we first specify some properties that human shape representations must have. These will help us to understand the central problem of bridging between early visual encoding (e.g., by local, orientation-sensitive units) and higher-level notions of shape. On one hand, it is clear that human perceptual abilities to see shape and shape similarities implicate more abstract symbolic coding than can be accomplished by sets of local orientations. On the other hand, it seems doubtful that our brains represent shapes with the arbitrary level of precision possible with mathematical techniques common in computer vision. For example, for the shapes of occluded contours, it has been argued [11] that experimental data is fit best by quintic (5th order) polynomials. This is no doubt a faithful description of curve-fitting results; however, one may wonder whether we should take seriously the idea that the brain really uses such a complex representation for shape and how it might generate quintic polynomials. Certainly no brain mechanism for generating them has yet been proposed. Moreover, such a representation, in a given case, would suggest a highly precise contour description, whereas psychophysical tests on human representations would likely show that our shape memory, at least, is substantially more vague.

For simplicity, we focus primarily, but not exclusively, on contour shape [12, 27]. Even the relatively basic shape notions we will consider evoke the issues of abstraction and simplification in shape representations that we wish to illustrate. Figure 18.10 illustrates some properties that seem to characterize human contour shape representations.

A shape representation for the contour given in Fig. 18.10A is sufficient to allow a shape match with one or more of the shapes in Figs. 18.10B, C, and D. This is possible despite changes in size, orientation, or even the elements comprising the figure. In terms of the lines or elements making up the figure, the display in 18.10C is most like 18.10A, yet inspection readily reveals that it is the only figure whose overall shape is different from 18.10A. Some of the key properties indicated by these simple shape-matching capabilities are that shape representations have some degree of scale invariance, orientation invariance, and that a common shape can be extracted despite differing constituent elements. These points were made long ago by the Gestalt psychologists (e.g., [34]), who emphasized that the simple summation

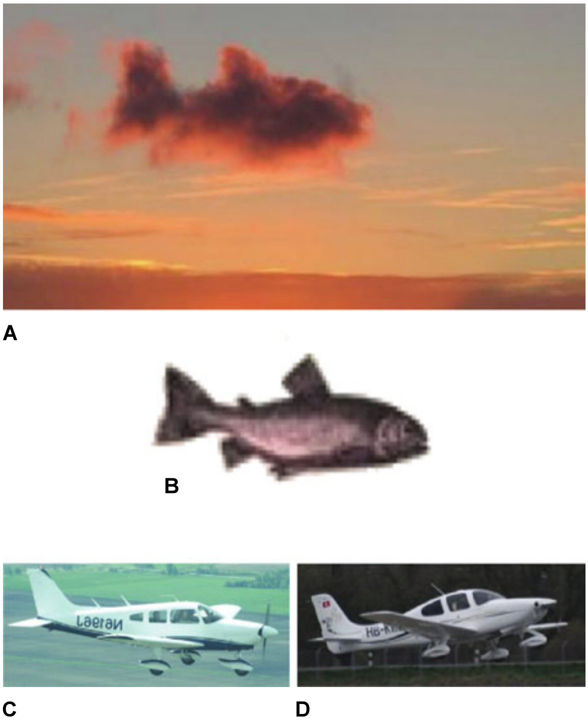


Fig. 18.11 Illustration of gist and similarity relations in shape representation. (See text.) (From http://lolyard.com/3448/cloud-fish)

of sensory elements did not comprise form; indeed, form consists of relations, which can be conveyed by many different kinds of sensory elements.

The modern version of the Gestalt point is fully relevant to primary issues in understanding shape and the connection between early visual coding and symbolic representations. We know that the patterns in Fig. 18.10 stimulate sets of spatially localized, orientation-sensitive units in the visual cortex (in V1 and V2). Yet human shape-matching performance clearly indicates that seeing the same or similar shape is not a matter of activating the same local orientation-sensitive units. The transformation of size changes the spatial frequency of the relevant units; changing orientation of the shape alters the relevant local orientations that are detected; and various elements can be used such that there is little or no overlap in populations of basic detectors that are activated and lead to perception of the same shape. How do we get from the stage of local orientation encoding to more abstract percepts and representations of shape?

Figure 18.11 illustrates two other crucial properties of human shape perception and representation. One we might call "gist." The cloud shown in Fig. 18.11A has quite ragged edges, including various protrusions and "frayed edges" in various places. A fully precise contour representation that matched all of the visible boundary points (e.g., what one might get by doing a precise, higher-order polynomial fit) would give a very jagged and complicated boundary contour representation. It is doubtful that, after looking briefly at such an image, we possess any such fully detailed representation. As a thought experiment, if we showed observers such images, took them away, and then presented new images in which the perturbations along the edges had been moved or changed, it is unlikely that observers would be good at detecting these changes (for a more detailed example, see [12]). Our representations encode the overall shape at a level that is likely to be relevant to our functioning in the world. Encoding all of the little wisps and deviations along the edges in this case are unlikely to be of functional importance (although given specialized tasks, this could change).

The other property illustrated in Fig. 18.11 is closely related to gist. It is that our shape representations support similarity relations in a constrained but flexible manner. The shape similarity of the cloud and the fish are obvious. Slightly more demanding is the question of which of two aircraft more closely matches the cloud shape. Pretty clearly it is the aircraft in Fig. 18.11D (the one on the right). These shape matching feats are remarkable because the matching images are far from a match at the pixel level or at the level of sets of local orientation detectors activated by the two patterns.

In summary, the visual system must somehow get from early local encodings of oriented contrast to more global and abstract shape representations. These representations are unlikely to be precise polynomial approximations to detailed boundaries, but are likely to be simplifications in some way. And these simplifications are likely to be the very properties that allow approximate matching to similar forms that are by no means identical, either at the one extreme of activating the same population of oriented units or at the other extreme of matching a precise mathematical description of a bounding contour.

18.10 Constant Curvature Coding: An Example of a Bridge Between Subsymbolic and Symbolic Shape Coding

The foregoing discussion of requirements of human contour shape representation may be useful in indicating important constraints on theories of biological shape representation, but they also represent a set of daunting challenges. Much of the point of this discussion is that we do not currently have suitable theories of shape that meet these requirements. There is no doubt, of course, that mathematical approaches for specifying shape are flexible enough such that we could specify symbolic representations that meet the requirements, but that would leave open the question of how such representations are acquired from the initial encoding of visual information. We do not offer a comprehensive answer to these problems, but we propose a scheme that addresses some particular issues, and, more generally, offers an existence proof of how more symbolic tokens might be acquired from subsymbolic precursors.

18.11 Early Symbolic Encoding of Contours: Arclets

Based on considerations of simplicity, coding efficiency, and some existing psychophysical and neurophysiological data, we have developed a scheme that uses the simplicity of the circle as the link between low- and higher-level vision [12, 27]. We propose that neural circuits exist that combine small groups of oriented units that are linked by constant turning angles, e.g., they encode constant curvature segments (including zero curvature) of contour shape. We call these *arclets*. Any open contour (including a part of the bounding contour of an object) may be described in terms of segments of constant curvature. In recent work, we have proposed two computational models of how this encoding could work, with the models differing in the tradeoff between the load in terms of number of segments and the fidelity of getting a near exact match to a viewed contour [12]. We refer the reader to that work for details.

For present purposes, the more important point is how arclets can operate as a bridge between subsymbolic and symbolic encoding. In their application to interpolation, activation initiated by real contours spreads along restricted paths in a network of oriented units; these paths consist of arclets. Because of this restriction, there is a unique path of interpolation connecting any relatable edges [29, Appendix A]. In their application to shape coding, arclets are symbolic tokens that are activated by signals in chains of several oriented units. This allows a natural means of handing off the information encoded by local oriented units to higher-level shape representations.

The central idea is that an important basic level of abstract shape encoding consists of contour representations comprised of one or more constant curvature segments. These middle-level shape representations result from detectors that are activated by sets of oriented units in particular relations to each other (cf., [10]). As illustrated in Fig. 18.12, a given arclet is activated if a chain of oriented units forming a collinear or co-circular path are simultaneously activated. At the bottom of this figure is the viewed object. The object activates sets of oriented units (shown as Gabor patches) in early cortical areas. Arclet detectors respond to chains of these units having a constant angular relation (turn angle).

This is the locus of the transition from local, contrast-sensitive elements to the first symbolic representation. The activated arclet token contains three pieces of information: the scale (spatial frequency) of the oriented units, the turn angle relating them (20 deg in the example given), and the number of oriented units (encoding segment length). We assume some system of competition to find the best-fitting arclet for any segment, as arclets of different scales and turn angles may fit to differing degrees.

Different arclets code different curvatures. Activation of a single arclet indicates the presence of that curvature at a certain position and orientation. The encoding of a constant curvature segment extends along a contour until a transition zone, at which arclets of that curvature exceed some threshold of accurately matching the contour (or are less well activated than some arclet having a different curvature value). A shape representation consists of a set of constant curvature values characterizing

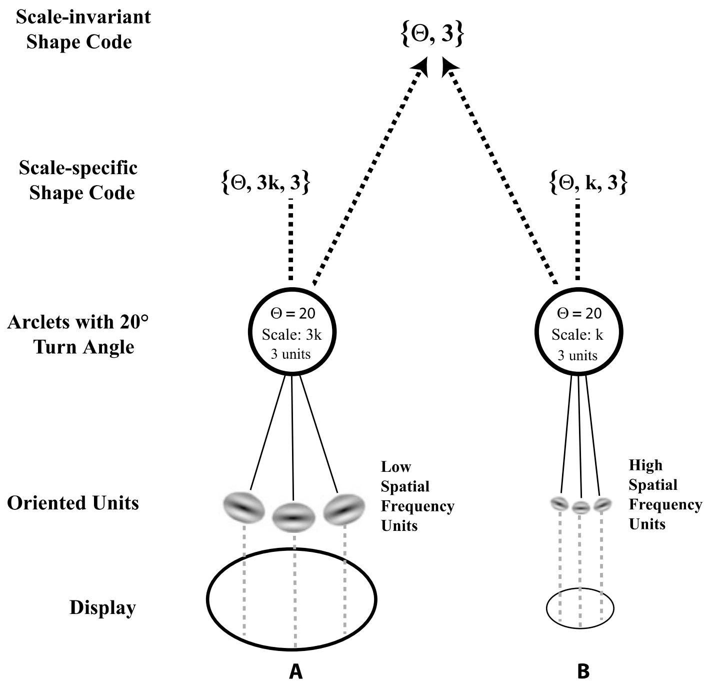


Fig. 18.12 Illustration of constant curvature segment encoding. (See text.) (From [27]. In M.A. Peterson, B. Gillam, & H.A. Sedgwick, (Eds.) *In the Mind's Eye: Julian Hochberg on the Perception of Pictures, Film, and the World*. New York: Oxford University Press. Reprinted with permission)

segments along a contour, along with some marking of transition zones between constant curvature segments. (For working models of this scheme, see [12]).

As shown in Fig. 18.12, arclets have the interesting property of permitting concurrent scale-variant and scale-invariant coding of contours. A problem for understanding invariance in human perception is that standard mathematical notions of curvature do not capture shape invariance. A large circle and a small circle obviously have the same shape, but they have very different curvatures (where curvature is given by the change in contour orientation per unit arc length). Typically, use of relative curvatures or normalization by some overall object size measurement is used to compare shapes in computer vision and some biological vision work [8, 22].

Arclets offer a means of achieving scale invariance in a more natural way. Because orientation-sensitive units in early visual areas exist across a range of spatial scales, arclets would similarly span this range. An interesting invariant characterizes arclets made of differently sized elements that are related by the same turn angle. As long as all elements within each arclet are of equal size, all arclets based on the same turn angle between oriented elements represent the *same scale-invariant shape*, i.e., shape pieces that differ only by a scalar. This is shown in Fig. 18.12, in which two arclets at different scales both encode segment lengths including the same number of oriented units in a chain having the same turn angle. Activating a best-fitting arclet at any scale therefore signals a unique number (based on the turn angle) that specifies scale-invariant shape for that part of the contour. Two circles of different sizes, for example, will have contours that best match arclets at different scales, but both arclets will have the same turn angle.

Remarkably, this property of obtaining size invariance for free comes from the use of oriented segments of finite lengths to encode curvature. Mathematically, a perfect description of a curve would have infinitesimal segments; orientation is constantly *changing* along a curve! Approximating curvature using units sensitive to orientations that are constant along their lengths would seem a necessary but regrettable compromise in encoding. It is, however, this characteristic that allows a scale-invariant curvature property to emerge automatically. Our analysis is consistent with the fact that the size of oriented units in human vision co-varies with their spatial frequency.

In Fig. 18.12, the two ellipses, having the same shape but different sizes, each have a constant curvature segment that is shown as encoded by arclets with the same turn angle and the same number of participating units (length). At the level of a scale-specific representation (allowing us to see that the two ellipses are different in size), the scale-specific arclet representation preserves the turn angle information and the scale of the elements in the best fitting arclet. At this level, the large ellipse is shown as having scale $3k$ and the smaller ellipse as having scale k. The scale invariant representation, in which the shape of the corresponding segment of each ellipse is encoded identically, simply drops out the scale term. Curved segments having the same turn angle and comprised of the same number of units specify the same perceived shape (relative curvature).

Because the arclets are encoding change information (turn angles), orientation invariance also comes naturally with this form of representation. Orientation invariance has limits in human form perception [43]. Analogous to the concurrent scale-variant and scale-invariant encoding, absolute orientation information of segments is likely preserved for some purposes, including form coding that has privileged reference axes.

Garrigan & Kellman [12] discuss alternative versions of an arclet-based code that trades off between complexity (in terms of number of parts) and fidelity (in terms of how faithfully the code represents the contour). Most contours in the world do not consist of constant curvature segments (as is true of the ellipse in Fig. 18.12), but they could be approximated to any level of precision by many small constant curvature pieces. A simpler code in terms of constant curvature segments would have fewer segments but more distortion. It seems likely that the precision of contour coding varies with attention and task demands.

There are alternative possibilities for early symbolic encoding of contours. Codes that utilize more complicated primitives, e.g., any spline fitting model, will outperform the arclet-based approach in some cases, but also have a number of shortcomings when considered as a model of contour shape representation that can handle the shape-related problems the human visual system encounters under normal viewing conditions. Consider, e.g., recognizing that one contour segment is part of another contour. More simple shape primitives (like the arclets) are less sensitive to long-range relationships along the contour, a characteristic that may be critical for matching the representation of a smaller contour segment to part of the representation of a larger contour.

The simplicity of the arclet representation is also an advantage as the problem of shape representation is scaled up to more ecological shapes. Contour shape representation likely precedes intermediate representations (e.g., surface shape) and the representation of the shapes of behaviorally important objects that may have additional complexities (e.g., an animal with articulating parts). A simpler code that does not leverage some of the more complicated, perhaps distal relationships among the features of a contour may be more robust when these additional complexities are included. Consider the problem of articulating parts. A shape code that represents the bounding contour of an object with a very small set of relatively complicated contour shape primitives will have little relationship to the representation of the bounding contour of that same object if one part of the object unexpectedly moves. In sum, besides the tradeoff between fidelity and complexity, there is likely a tradeoff between efficiency and stability. The arclets are not the most efficient representation of contour shape, but they may lead to a more stable representation than more sophisticated primitives that leverage regularities that do not persist across viewing conditions.

18.12 Evidence for Constant Curvature Coding in Human Shape Perception

The arclets approach to constant curvature encoding of contours offers an example of how subsymbolic encoding might lead to more abstract shape codes. This specific proposal of constant curvature coding is also consistent with a variety of evidence in human vision, including results of recent research.

Pizlo, Salach-Golyska, & Rosenfeld [41] compared detection performance for a curve formed from dots arranged in straight lines, dots arranged in circular arcs, and dots arranged in various types of irregular paths. They found that straight lines were easiest to detect, but that circular arcs were easier to detect than irregular paths (provided the change in curvature along the irregular path was not too small). Pizlo, et al. also found that circular arcs were significantly easier than all the irregular paths they tested when the subject was given prior information about the shape of the target. These results are consistent with the importance of constant curvature extraction and memory in shape perception.

More recently, Achtman, Hess, & Wang [1] used a Gabor-path detection paradigm and showed that circular paths were more easily detected than radial or spiral paths. Similar detection threshold advantages for circles have been found using Glass patterns [35, 45, 53]. Other evidence, however, suggests that the primitives for form perception may include both circular and spiral pooling mechanisms [52].

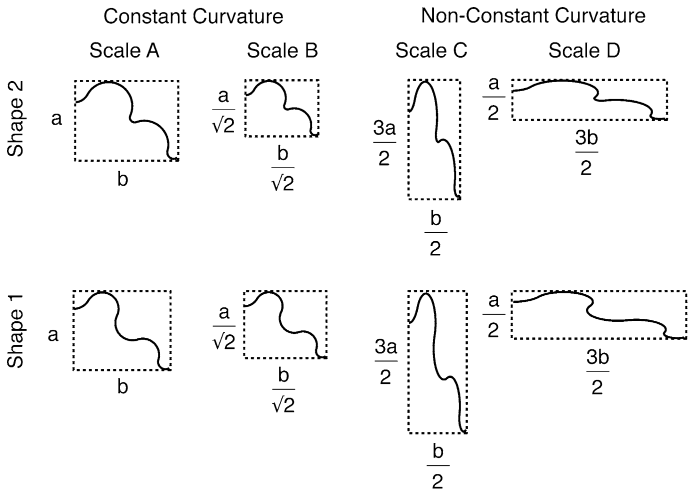


Fig. 18.13 Stimuli in experiments on constant curvature segment coding. Shape 1 is composed of five circular segments with differing radii. Each segment has constant curvature. Scaling horizontally and vertically by the same amount preserved the constant curvature (scale B), whereas scaling along the two dimensions by different amounts produces regions of non-constant curvature (scales C and D). These are all "matching" shapes. Shape 2 is created by changing the curvature of one of the circular segments of Shape 1 and is a "non-match" shape. (From P. Garrigan and P.J. Kellman, 2011, *Perception*, *40*(11), p. 1297. Reprinted with permission)

Neurophysiological evidence also supports a special role for constant-curvature encoding in shape perception. Single cell recordings in macaque monkeys are consistent with the idea that intermediate visual areas such as V4 may be representing object-oriented contour curvature [38–40]. These investigators suggest that representations of overall shapes can be derived from the collective output of such cells.

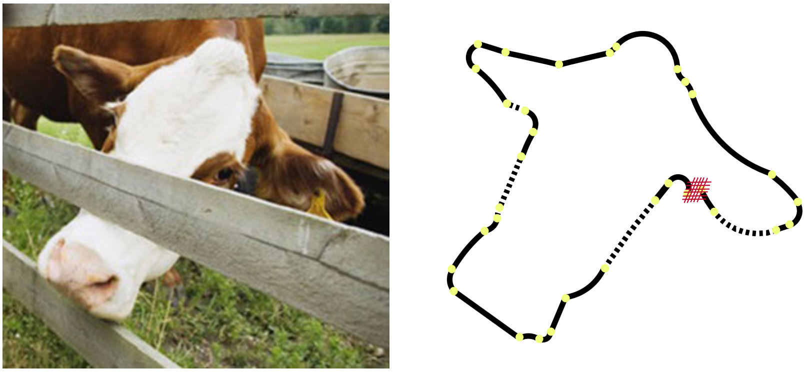
In recent psychophysical work, Garrigan and Kellman [12] used open contours to investigate the role of constant curvature in shape representations. Subjects judged whether two sequentially presented contour segments were the same or not, allowing for scale, rotation and translation transformations. The stimuli were created by combining five circular segments of differing radii and spans (Fig. 18.13). Because they were constructed from circles, each segment had a constant curvature. Scaling the shape by an equal amount horizontally and vertically preserved the constant curvature, while scaling by different amounts along each dimension produced non-constant curvature segments. Non-matching shapes were created by changing the curvature of one of the circular segments (see Shape 2, Fig. 18.13).

Subjects were reliably more accurate in matching constant curvature shapes than non-constant curvature shapes. Even when all transformations were removed so that the two stimuli were exactly identical, subjects were more accurate in matching constant curvature shapes, when shapes had to be compared across a retention interval of 1000 ms. Similar recognition performance was observed for both shape types, however, when they were compared at the same size and viewpoint and the retention interval was reduced to 500 ms. These findings are consistent with a symbolic encoding of 2-D contour shapes into constant curvature parts when the retention intervals over which shapes must be stored exceeds the duration of initial, transient, visual representations.

These experiments and the arclets model provide a plausible proposal for the how a location-specific, subsymbolic representation might transition to an abstract, symbolic one. Local edge information may be integrated into a scale and rotation invariant representation of contour curvature. They represent modest steps, as efforts to understand abstraction in perceptual representations is a multifaceted, challenging, and ongoing effort. These proposals do, however, offer an existence proof related to some of the most open-ended questions in understanding perceptual representations: How does the visual system construct abstract, flexible, functionally useful shape representations from the early encoding of local, literal image properties? Constant curvature representations of contours are computationally possible, and consistent with both properties of early cortical units in vision and some results suggesting curvature coding in visual area V4.

18.13 Connecting Contour Interpolation and Shape Descriptions

Consistent with the complexity of problems of contour, object, and shape perception, even this short overview has covered a lot of ground. We noted the relevance of interpolation processes to shape, in that shape descriptions typically encompass not image fragments, but the outputs of object formation processes. We then focused on representations of contour shape, which the visual system may obtain in symbolic form by recoding object contours in terms of constant curvature parts. Figure 18.14 provides an example pulling together these themes, using the picture of a cow from Fig. 18.3. The cow's partially occluded head in the original image (Fig. 18.14A) is shown represented as a completed, constant-curvature based shape representation (Fig. 18.14B). Primary edges were found using a version of the Canny edge detector [7]. (This approach was used for simplicity here, although its outputs are highly consistent with some models that utilize neutrally plausible local units to do initial edge finding (e.g., [23]).) In Fig. 18.14B, the edges of the cow's head are shown after recoding as constant curvature segments, consistent with the contour curvature model of Garrigan & Kellman [12]. Dotted lines indicate amodally completed contours; these have been interpolated following the rules of contour relatability, with the resulting interpolated contours also represented with constant curvature parts. The interpolation model and the contour shape model are guaranteed to be consistent, as any interpolated contour consistent with relatability can be described uniquely as consisting of one zero curvature and one constant curvature segment ([29], cf., [51]).



A B

Fig. 18.14 Example of interpolation and constant curvature coding in shape perception. The cow's partially occluded head in the original image (**A**) is shown represented as a completed, constant-curvature based shape representation (**B**). Primary edges were obtained from the raw image using a common edge-detection operator. Edges of the cow's head in the image on the right have been recoded as constant curvature segments, consistent with the contour curvature model of Garrigan & Kellman [12]. *Dotted lines* indicate amodally completed contours following the rules of contour relatability, with the resulting interpolated contours also represented with constant curvature parts. The *cross-hatched area in the display on the right* indicates an area where edges and textures are ambiguous and do not permit clear interpolation. (See text)

This example is not meant to minimize challenging problems that remain in understanding the transition from subsymbolic to symbolic coding, even in the relatively simple domain of contour perception. As we have discussed elsewhere, although some extant interpolation models can use raw images as their input, they will be improved when certain symbolic encoding is added, such as representing a unique edge orientation at each contour junction rather than a distribution of orientation activations (for discussion, see Kalar et al. [23]). Natural scenes may also have areas for which the outputs of edge finding and/or interpolation models are indeterminate, as in the cross-hatched area indicated in Fig. 18.14B. Sometimes these outputs are likely consistent with some indeterminacy in actual perception, but in other cases they likely indicate limitations of current models. Regarding the recoding of contours into segments of constant curvature, the model of Garrigan & Kellman [12] is a working algorithm that takes a contour specified in terms of local orientation values and produces constant curvature segments as outputs, but no full implementation yet exists in terms of attaining each local orientation value from the outputs of separate, local, orientation-sensitive units at multiple scales. Moreover, versions of the model vary in their tradeoff of fidelity (minimizing differences from the input image, but requiring greater numbers of segments in the approximation) and economy (accepting limits in fidelity due to some capacity or complexity limit). We have proposed that the visual system may similarly adjust contour shape coding for greater fidelity or greater economy [12], depending on task demands and attention, but the specifics are not known.

18.14 Summary

Shape perception and representation pose fascinating challenges in vision science. In this article, we have focused on perhaps the greatest theoretical chasm in understanding shape: the origin of abstract, symbolic representations. Perceived shape is not a readout of image characteristics, nor is it a collection of activations of early orientation-sensitive units. Image regions do not receive shape descriptions in human perception; rather, the shapes we record relate to objects formed by interpolation processes that may connect various separated regions. Both the shapes of interpolated contours and of real contours share representational formats that make possible invariant shape recognition despite certain scale and orientation changes, matching of shapes despite different constituent elements, and extraction of gist in shape encoding, allowing detection of shape similarities. Understanding the nature of these symbolic representations, and how they are constructed from earlier encodings, is a complex task. Vision science is fortunate in having some understanding of initial subsymbolic encoding by neural units, and having also a number of middle or high-level vision models that begin with representations that are already abstract. The challenge is to discover how these meet in the middle—how we attain more global, symbolic, interpreted descriptions from local, non-symbolic encodings. For contour and object boundary representations, extraction of constant curvature segments as basic tokens of early symbolic representations may comprise an important step, one consistent with psychophysical and neurophysiological data, and one that illustrates how the visual system may approach "the middle game."

Acknowledgements We thank Brian Keane, Evan Palmer, and Hongjing Lu for helpful discussions and Rachel Older for general assistance. Portions of the research reported here were supported by National Eye Institute Grant EY13518 to PJK.

References

1. Achtman RL, Hess RF, Wang Y-Z (2003) Sensitivity for global shape detection. J Vis 3:616–624
2. Albert MK (2001) Surface perception and the generic view principle. Trends Cogn Sci 5(5):197–203
3. Albert MK, Hoffman DD (2000) The generic-viewpoint assumption and illusory contours. Perception 29(3):303–312
4. Albert MK, Tse P (2000) The role of surface attraction in perceiving volumetric shape. Perception 29:409–420
5. Banton T, Levi DM (1992) The perceived strength of illusory contours. Percept Psychophys 52(6):676–684
6. Baylis GC, Driver J (1993) Visual attention and objects: evidence for hierarchical coding of location. J Exp Psychol Hum Percept Perform 19(3):451–470
7. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 8(6):679–698
8. Costa LF, Cesar RM (2001) Shape analysis and classification: theory and practice. CRC Press, Boca Raton

9. Fantoni C, Hilger J, Gerbino W, Kellman PJ (2008) Surface interpolation and 3d relatability. J Vis 8(7):1–19
10. Field DJ, Hayes A, Hess RF (1993) Contour integration by the human visual system: evidence for a local "Association field". Vis Res 33(2):173–193
11. Fulvio JM, Singh M, Maloney LT (2009) An experimental criterion for consistency in interpolation of partly occluded contours. J Vis 9(4):1–19
12. Garrigan P, Kellman PJ (2011) The role of constant curvature in 2-d contour shape representations. Perception 40(11):1290–1308
13. Geisler WS, Perry JS (2009) Contour statistics in natural images: grouping across occlusions. Vis Neurosci 26:109–121
14. Geisler WS, Perry JS, Super BJ, Gallogly DP (2001) Edge co-occurrence in natural images predicts contour performance. Vis Res 41:711–724
15. Ghose T, Erlikhman G, Kellman PJ (2011) Spatiotemporal object formation: contour vs surface interpolation. Perception 40 ECVP Abstract Supplement, p 59
16. Gibson JJ (1979) The ecological approach to visual perception. Houghton Mifflin, Boston
17. Grossberg S, Mingolla E (1985) Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. Psychol Rev 92:173–211
18. Guttman SE, Kellman PJ (2004) Contour interpolation revealed by a dot localization paradigm. Vis Res 44:1799–1815
19. Guttman SE, Sekuler AB, Kellman PJ (2003) Temporal variations in visual completion: a reflection of spatial limits? J Exp Psychol Hum Percept Perform 29(6):1211–1227
20. Heitger F, Rosenthaler L, von der Heydt R, Peterhans E, Kübler O (1992) Simulation of neural contour mechanisms: from simple to end-stopped cells. Vis Res 32(5):963–981
21. Heitger F, von der Heydt R, Peterhans E, Rosenthaler L, Kübler O (1998) Simulation of neural contour mechanisms: representing anomalous contours. Image Vis Comput 16(6–7):407–421
22. Hoffman DD, Singh M (1997) Salience of visual parts. Cognition 63:29–78
23. Kalar DJ, Garrigan P, Wickens TD, Hilger JD, Kellman PJ (2010) A unified model of illusory and occluded contour interpolation. Vis Res 50:284–299
24. Keane BP, Lu H, Papathomas TV, Silverstein SM, Kellman PJ Reinterpreting behavioral receptive fields: lightness induction alters visually completed shape, accepted pending minor revision. PLoS ONE
25. Kellman PJ (2003) Interpolation processes in the visual perception of objects. Neural Netw 16:915–923
26. Kellman PJ (2003) Segmentation and grouping in object perception: a 4-dimensional approach. In: Behrmann M, Kimchi R (eds) Perceptual organization in vision: behavioral and neural perspectives: the 31st Carnegie symposium on cognition. Erlbaum, Hillsdale
27. Kellman PJ, Garrigan P (2007) Segmentation, grouping, and shape: some Hochbergian questions. In: Peterson MA, Gillam B, Sedgwick HA (eds) In the mind's eye: Julian Hochberg on the perception of pictures, film, and the world. Oxford University Press, New York
28. Kellman PJ, Loukides MG (1987) An object perception approach to static and kinetic subjective contours. In: Meyer G, Petry G (eds) The perception of illusory contours. Springer, New York, pp 151–164
29. Kellman PJ, Shipley TF (1991) A theory of visual interpolation in object perception. Cogn Psychol 23:141–221
30. Kellman PJ, Yin C, Shipley TF (1998) A common mechanism for illusory and occluded object completion. J Exp Psychol Hum Percept Perform 24(3):859–869
31. Kellman PJ, Guttman SE, Wickens TD (2001) Geometric and neural models of object completion. In: Shipley TF, Kellman PJ (eds) From fragments to objects: segmentation and grouping in vision. Elsevier, Oxford
32. Kellman PJ, Garrigan P, Shipley TF (2005) Object interpolation in three dimensions. Psychol Rev 112(3):586–609
33. Kellman PJ, Garrigan P, Shipley TF, Yin C, Machado L (2005) J Exp Psychol Hum Percept Perform 31(3):558–583
34. Koffka K (1935) Principles of Gestalt psychology. Harcourt Brace, New York

35. Kurki I, Saarinen J (2004) Shape perception in human vision: specialized detectors for concentric spatial structures? Neurosci Lett 360(1–2):100–102
36. Marr D (1982) Vision. Freeman, New York
37. Palmer EM, Kellman PJ, Shipley TF (2006) A theory of dynamic occluded and illusory object perception. J Exp Psychol Gen 135:513–541
38. Pasupathy A, Conner CE (1999) Responses of contour features in macaque area V4. J Neurophysiol 82:2490–2502
39. Pasupathy A, Conner CE (2001) Shape representation in area V4: position-specific tuning for boundary conformation. J Neurophysiol 86:2505–2519
40. Pasupathy A, Conner CE (2002) Population coding of shape in area V4. Nat Neurosci 5:1332–1338
41. Pizlo Z, Salach-Golyska M, Roenfeld A (1997) Curve detection in a noisy image. Vis Res 37(9):1217–1241
42. Ringach DL, Shapley R (1996) Spatial and temporal properties of illusory contours and amodal boundary completion. Vis Res 36:3037–3050
43. Rock I (1973) Orientation and form. Academic Press, San Diego
44. Rubin N (2001) The role of junctions in surface completion and contour matching. Perception 30:339–366
45. Seu L, Ferrera VP (2001) Detection thresholds for spiral glass patters. Vis Res 41(28):3785–3790
46. Shipley TF, Kellman PJ (1990) The role of discontinuities in the perception of subjective contours. Percept Psychophys 48(3):259–270
47. Shipley TF, Kellman PJ (1992) Perception of partly occluded objects and illusory figures: evidence for an identity hypothesis. J Exp Psychol Hum Percept Perform 18(1):106–120
48. Shipley TF, Kellman PJ (1992) Strength of visual interpolation depends on the ratio of physically-specified to total edge length. Percept Psychophys 52(1):97–106
49. Singh M, Hoffman DD (1999) Completing visual contours: the relationship between relatability and minimizing inflections. Percept Psychophys 61(5):943–951
50. Singh M, Hoffman DD, Albert MK (1999) Contour completion and relative depth: Petter's rule and support ratio. Psychol Sci 10:423–428
51. Ullmann S (1979) The shape of subjective contours and a model for their generation. Biol Cybern
52. Webb BS, Roach NW, Peirce JW (2008) Masking exposes multiple global form mechanisms. J Vis 8(9):1–10
53. Wilson HR, Wilkinson F, Asaad W (1997) Concentric orientation summation in human form vision. Vis Res 37(17):2325–2330
54. Yin C, Kellman PJ, Shipley TF (1997) Surface completion complements boundary interpolation in the visual integration of partly occluded objects. Perception 26:1459–1479
55. Yin C, Kellman PJ, Shipley TF (2000) Surface integration influences depth discrimination. Vis Res 40:1969–1978
56. Zemel R, Behrmann M, Mozer M, Bavelier D (2002) Experience-dependent perceptual grouping and object-based attention. J Exp Psychol Hum Percept Perform 28:202–217